Building a database of Alzheimer's patients

Jicheng Lu

1 Introduction

In this report, we introduce the database of Alzheimer's patients. The entire database consists of four sections: Demographics, Medical Test, Medical History and Personal Characteristics. Each section contains several related medical or physical attributes (Details will be shown below). We first present the method to generate the data and describe the attributes in each section. Then we show the generated data in MySQL database. Finally, we point out some issues in the current database.

2 Method

For the demographics section, we generate the data using the Bayes formula. In doing so, we can get the diagnosis for each patient based on his or her basic information, such as sex, age, year of education, race and hispanic. Then based on the result, we can generally split the data into two categories: normal and abnormal.

For a normal subject, we assume the subject has no disease or symptom, doesn't take any medicine or drug, all the medical test results are normal. In addition, some indexes that represent living habits, such as BMI, coffee or smoke consumption, are randomly sampled.

For those mentally abnormal subjects, we randomly assign related items for each attribute when processing Medical Test, Medical History and Personal Characteristics.

3 Database Description

In this section, we present the attributes in each table as well as the items in each attribute.

3.1 Demographics

The Demographics table contains the basic information of patients, such as age, sex, year of education, year of work, race and hispanic. We also assign the diagnosis to each patient, including "Normal", "Impaired, not MCI", "MCI", "Dementia". The distributions of the basic information and diagnosis are obtained from NACC (National Alzheimer's Coordinating Center, https://www.alz.washington.edu/WEB/UDSonepage.pdf).

The items in each attribute are shown in Fig.1.

```
attributes = {}
attributes["age"] = ["<65", "65-84", ">=85"]
attributes["yearOfEducation"] = ["<=12", "13-16", ">=17"]
attributes["yearOfWork"] = ["<=12", "13-16", ">=17"]
attributes["sex"] = ["Male", "Female"]
attributes["race"] = ["White", "Black or African American", \
                      "American Indian or Alaska Native", "Native Hawaiian", \
                      "Asian", "Multiracial", "Unknown"]
attributes["hispanic"] = ["No", "Yes"]
```

Figure 1: Items in each attribute of Demographics.

3.2 Medical Test

The Medical Test table contains two aspects:

- Medical scores, such as MMSE (Mini Mental State Examination), HIS (Hachinski ischemic score), etc.
- Result of certain medical test, such as CT showing absence of clinically significant focal lesion, etc.

The attributes in the Medical Test table are shown in Fig.2.

```
items = ["MRI", "CT", "urineTest", "petImaging", "clinicalDementiaRating",\
        "MMSE", "modHachinskiScore", "hamDepressRatingScaleScore", "frontalSystemsBehaviorScale", \
        "BMI", "restingPulse", "systolicBloodPressure", "diastolicBloodPressure"]
```

Figure 2: Attributes of Medical Test.

Since we do not have the distribution data, the probabilities of each item in this table are self-determined. The items in each attribute are shown in Fig.3.

```
attributes = {}
attributes["MRI"] = ["consistent ad", "intervening neurological disease", \
                     "cerebral microhemorrhages", "infection", "infarction", \
                     "fluid attenuation inversion recovery", "focal lesion", \
                     "other abnormality", "normal"]
attributes["CT"] = ["consistent ad", "intervening neurological disease", \
                    "infection", "infarction", \
                    "focal lesion", "other abnormality", "normal"]
attributes["urineTest"] = ["drug abuse", "normal"]
attributes["petImaging"] = ["amyloid burden with ad", "normal"]
attributes["clinicalDementiaRating"] = ["0", "0.5", "1", "2", "3"]
attributes["MMSE"] = ["0-12", "13-20", "21-24", "25-30"]
attributes["modHachinskiScore"] = ["0-4", "5-18"]
attributes["hamDepressRatingScaleScore"] = ["0-9", "10-13", "14-17", "18-54"]
attributes["frontalSystemsBehaviorScale"] = ["14-28", "29-52", "53-70"] # only for apathy - 14 items
attributes["BMI"] = ["10-18.5", "19-25", "26-30", "31-40"]
attributes["restingPulse"] = ["30-59", "60-80", "81-100"]
attributes["systolicBloodPressure"] = ["100-119", "120-129", "130-139", "140-179", "180-210"]
attributes["diastolicBloodPressure"] = ["60-80", "60-80", "80-89", "90-119", "120-180"]
```

Figure 3: Items in each attribute of Medical Test.

3.3 Medical History

The Medical History contains three aspects:

- Whether a patient has a certain disease or symptom, such as psychiatric disorders, seizures, etc.
- Whether a patient has a certain habit in a period of time, such as smoking, alcohol, etc.
- Whether a patient uses a certain medicine in a period of time, such as phenytoin, carbamazepine, etc.

The attributes in the Medical History table are shown in Fig.4.

```
items = ["useMedicine", "useDuration", "symptom", "symDuration", "symCondition", \
        "coffee", "smoke", "alcoholAbuse", "alcoholDep", "drugAbuse", "drugDep"]
```

Figure 4: Attributes of Medical History.

Before showing the attributes in the medical history table, we summarized the types of medicines mentioned in the criteria. Currently, there are 37 types of medicines (Fig.5). The medicine is stored in the Medicine table, where each medicine is assigned with a unique ID. The "useMedicine" attribute in the Medical History table corresponds to each ID in the Medicine table.

```
medicines = ["galantamine", "rivastigmine", "donepezil", "tacrine", "analgesics", \
             "sinemet", "amantadine", "bromocriptine", "pergolide", \
             "selegiline", "estrogen", "vitaminE", "memantine", \
             "antidepressant", "ginkgo biloba", "insulin", "betablockers", \
             "narcotics", "methyldopa", "clonidine", "antiparkinsonian", \
             "bromocriptine", "neuroleptics", "benzodiazepines", "barbituates", \
             "anxiolytics", "sedative hypnotics", "corticosteroids", "anticonvulsants", \
             "phenytoin", "phenobarbital", "carbamazepine", "warfarin", "vitamin", \
             "anticoagulant", "cholinesterase inhibitor", "other", "none"]
```

Figure 5: Items in Medicine.

Next, we present the items of each attribute in medical history , Fig.6. The probabilities of each item of attributes in Medical History are self-determined because of lack of data. **Note** that the "symptom" and "useMedicine" are not necessarily related with each other. That is, a medicine that a patient takes may not cure for the symptom that the patient has. Moreover, for each sick patient, he may **at most** have three different symptoms and use three different medicines.

```
attributes = {}
attributes["useMedicine"] = [str(i) for i in range(1, num_med)]
attributes["useDuration"] = ["1m", "2m", "6m", "1y", "2y", "3y"] # less than
attributes["symptom"] = ["head trauma", "peptic ulcer", "arrhythmia", "cancer", \
                         "malignant melanoma", "stroke", "epilepsy", \
                         "blindness", "deafness", "hepatic", "renal", \
                         "seizures", "mental disorder", "schizophrenia", "memory difficulty", \
                         "asthma", "diabetes", "hypertension", "coagulopathy", \
                         "gastrointestinal", "cerebrovascular", "pulmonary", "autoimmune", \
                         "cardiovascular", "other"]
attributes["symDuration"] = ["1m", "2m", "6m", "1y", "2y", "3y"] # less than
attributes["symCondition"] = ["mild", "moderate", "severe"]
attributes["coffee"] = ["0-5", "6-10"] # cups per day
attributes["smoke"] = ["0-10", "11-20", "21-30"]
attributes["alcoholAbuse"] = ["No", "Yes"]
attributes["alcoholDep"] = ["No", "Yes"]
attributes["drugAbuse"] = ["No", "Yes"]
attributes["drugDep"] = ["No", "Yes"]
```

Figure 6: Items in each attribute of Medical History.

3.4 Personal Characteristics

For the personal characteristics, it mainly contains the willingness, physical ability, caregiver condition and other features:

- Personal willingness, such as willingness to visit places, etc.
- Personal ability, such as normal ability to speak or understand, etc.
- The characteristics of patient's caregiver, such as willingness to accompany the patient, etc.
- Other features: disability, presence of metal devices inside body, etc.

The attributes in the Personal Characteristics table are shown in Fig.7.

```
items = ["speaking", "understand", "ingestOralMed", "resideInNurseFacility", \
         "hasCareGiver", "hasInformant", "isOutpatient", "willingness", "agreeMed", \
         "hasDisability", "medicalCondition", "presenceOfMetal", "isInOtherTrial", "comfortWithNet"]
```

Figure 7: Attributes of Personal Characteristics.

Again, the probabilities of each item in this table are self-determined. The items in each attribute are shown in Fig.8.

```
attributes = {}
attributes["speaking"] = ["Bad", "Moderate", "Good"]
attributes["understand"] = ["Bad", "Moderate", "Good"]
attributes["ingestOralMed"] = ["No", "Yes"]
attributes["resideInNurseFacility"] = ["No", "Yes"]
attributes["hasCareGiver"] = ["No", "Yes but NA", "Yes and available"]
attributes["hasInformant"] = ["No", "Yes but NA", "Yes and available"]
attributes["isOutpatient"] = ["No", "Yes"]
attributes["willingness"] = ["No", "Yes"]
attributes["agreeMed"] = ["No", "Yes"]
attributes["hasDisability"] = ["No", "Moderate", "Severe"]
attributes["medicalCondition"] = ["Bad", "Moderate", "Good"]
attributes["presenceOfMetal"] = ["No", "Yes"]
attributes["isInOtherTrial"] = ["No", "Yes"]
attributes["comfortWithNet"] = ["No", "Yes"]
```

Figure 8: Items in each attribute of Personal Characteristics.

4 Generated data

Here we randomly generate 10 patients data in the four tables. The results are shown in the following:

id	age	yearOfEducation	yearOfWork	sex	race	hispanic	diagnosis
1	78	23	16	Male	White	No	Dementia
2	81	22	14	Female	White	No	Normal
3	81	22	14	Female	White	No	Normal
4	71	9	13	Male	White	No	Dementia
5	66	11	11	Female	White	No	Dementia
6	91	17	11	Female	Black or African American	No	Dementia
7	65	13	18	Male	White	No	Dementia
8	76	22	20	Male	White	No	Normal
9	72	10	13	Male	White	No	Dementia
10	44	13	21	Female	White	No	Normal

Figure 9: Demographics table.

id	patientId	MRI	CT	urineTest	petImaging	clinicalDementiaRating	MMSE	modHachinskiScore	hamDepressRatingScaleScore
1	1	cerebral microhemorrhages	normal	normal	normal	2	15	2	11
2	2	normal	normal	normal	normal	0	30	0	0
3	3	normal	normal	normal	normal	0	30	0	0
4	4	consistent ad	intervening neurological disease	normal	normal	0.5	23	13	3
5	5	consistent ad	intervening neurological disease	normal	amyloid burden with ad	0.5	28	3	14
6	6	fluid attenuation inversion recovery	other abnormality	normal	amyloid burden with ad	1	21	13	1
7	7	infarction	consistent ad	normal	amyloid burden with ad	2	25	13	11
8	8	normal	normal	normal	normal	0	30	0	0
9	9	other abnormality	normal	normal	amyloid burden with ad	0.5	21	15	4
10	10	normal	normal	normal	normal	0	30	0	0

Figure 10: Medical Test table (part 1).

frontalSystemsBehaviorScale	BMI	restingPulse	systolicBloodPressure	diastolicBloodPressure
16	16.6	78	123	78
14	10.2	73	158	108
14	26.6	72	185	167
54	20.0	95	196	144
59	28.0	82	137	80
26	23.0	91	121	71
22	20.4	69	136	84
14	26.0	63	136	87
38	23.8	35	134	82
14	17.0	71	166	113

Figure 11: Medical Test table (part 2).

id	patientId	useMedicine	useDuration	symptom	symDuration	symCondition	coffee	smoke	alcoholAbuse	alcoholDep	drugAbuse	drugDep
1	1	10	1m	head trauma	6m	moderate	2	0	No	No	No	No
2	1	32	2y	other	3y	moderate	2	0	No	No	No	No
3	2	38	0	none	0	mild	9	29	No	No	No	No
4	3	38	0	none	0	mild	2	25	No	No	No	No
5	4	15	3y	epilepsy	1y	mild	1	0	No	No	No	No
6	4	30	2m	blindness	2y	mild	1	0	No	No	No	No
7	4	1	2y	asthma	3y	mild	1	0	No	No	No	No
8	5	37	1m	memory difficulty	6m	mild	1	3	No	Yes	No	No
9	5	1	1y	head trauma	2m	mild	1	3	No	Yes	No	No
10	6	19	6m	head trauma	2y	mild	6	15	No	No	No	No
11	7	14	6m	head trauma	3y	mild	0	24	No	No	No	No
12	8	38	0	none	0	mild	6	27	No	No	No	No
13	9	21	2y	coagulopathy	1m	severe	0	19	No	No	No	No
14	9	9	6m	autoimmune	2y	severe	0	19	No	No	No	No
15	10	38	0	none	0	mild	3	2	No	No	No	No

Figure 12: Medical History table.

id	patientId	speaking	understand	ingestOralMed	residelnNurseFacility	hasCareGiver	hasInformant	isOutpatient	willingness
1	1	Good	Good	Yes	Yes	Yes and available	Yes and available	Yes	Yes
2	2	Good	Good	Yes	No	No	No	No	No
3	3	Good	Good	Yes	No	No	No	No	No
4	4	Moderate	Good	Yes	No	No	Yes and available	No	Yes
5	5	Moderate	Moderate	Yes	No	No	Yes and available	Yes	Yes
6	6	Moderate	Moderate	Yes	No	No	Yes and available	Yes	Yes
7	7	Good	Good	Yes	Yes	Yes and available	Yes but NA	Yes	Yes
8	8	Good	Good	Yes	No	No	No	No	No
9	9	Good	Good	Yes	No	Yes and available	Yes and available	No	Yes
10	10	Good	Good	Yes	No	No	No	No	No

Figure 13: Personal Characteristics table (part 1).

agreeMed	hasDisability	medicalCondition	presenceOfMetal	isInOtherTrial	comfortWithNet
Yes	No	Good	No	No	Yes
No	No	Good	No	No	Yes
No	No	Good	No	No	Yes
No	No	Good	No	No	Yes
Yes	Moderate	Good	No	No	Yes
Yes	Moderate	Moderate	No	No	Yes
Yes	Moderate	Moderate	No	No	Yes
No	No	Good	No	No	Yes
Yes	No	Good	No	No	Yes
No	No	Good	No	No	Yes

Figure 14: Personal Characteristics table (part 2).

5 Issues

There are several issues in the current database:

- The probability distributions of items in Medical Test, Medical History and Personal Characteristics are unavailable so far.

- The symptom and medicine are not necessarily consistent in Medical History.
- Some items may need to be adjusted when the medical data is available.